DOI：10.14005/j.cnki.issn1672-7673.20230719.001

BFOSC 长缝光谱自动处理软件

宋得阳[1,2,3]，张君波[4,5]，王 靓[1,2,3]*

（1. 中国科学院南京天文光学技术研究所，江苏 南京 210042；2. 中国科学院天文光学技术重点实验室（南京天文光学技术研究所），江苏 南京 210042；3. 中国科学院大学，北京 100049；4. 中国科学院国家天文台，北京 100101；5. 中国科学院光学天文重点实验室（国家天文台），北京 100101）

摘要：在天文观测中，长缝光谱仪是一种常用的光谱观测设备。国家天文台兴隆观测基地2.16 m望远镜搭载的北京暗天体摄谱仪(Beijing Faint Object Spectrograph and Camera, BFOSC)的长缝光谱模式应用广泛，但没有配备专用的数据处理软件，通常用户观测后采用IRAF(Image Reduction and Analysis Facility)软件利用交互方式处理。介绍了一款针对BFOSC长缝光谱的自动化数据处理软件，数据处理过程包括本底(bias)和平场(flat)改正、背景改正、抽取光谱和波长定标等，展示了对实测数据进行处理的结果，并与IRAF软件处理的结果进行了对比。该软件在完成所有光谱处理必需步骤的基础上实现了自动化运行，并采用优化抽谱方法，能够自动扣除大部分宇宙线，保证了高精度的输出结果。

关键词：长缝光谱；优化抽谱方法；自动处理程序

中图分类号：TP391.41；P114 **文献标识码**：A **文章编号**：1672-7673(2023)06-0591-09

我国自主研制的第1台2 m级望远镜——国家天文台兴隆基地的2.16 m望远镜，配备了北京暗天体摄谱仪[1]。这台仪器具备5种不同的观测模式，其中长缝光谱是应用最为广泛的一种模式，可以在不同的缝宽、光栅和滤光片组合之间进行切换。2016年更换CCD相机后最新的色散、波长范围等参数如表1[1-3]。据公开资料统计，2021年兴隆观测基地为2.16 m望远镜用户分配的252个观测夜中，使用BFOSC的有165夜，占比超过65%。

BFOSC测光模式下的数据可以使用文[4]开发的程序处理。但长缝光谱部分尚没有专门的数据处理软件，观测者一般自行使用IRAF[5-6]等软件利用交互方式对原始光谱图像进行处理，获得一维谱。但这类软件并不是针对BFOSC特点设计的[7]，没有针对长缝光谱的特点进行优化，且美国国立光学天文台(National Optical Astronomy Observatory, NOAO)已经宣布自2013年起不再对IRAF进行官方支持。

本文针对BFOSC长缝光谱，开发了一套光谱处理软件，实现了自动化处理光谱数据。该程序可以读取用户指定的光谱仪原始数据，自动进行图像合并、波长定标、目标源光谱抽取等步骤，在关键步骤中保存中间结果并进行可视化。

表1 BFOSC不同光栅及基本性能参数

Table 1 Grating parameters of BFOSC

Name	Rec. Lin. Disp. /(nm·mm^{-1})	Disp. /(nm·pix^{-1})	Wav. Range /nm
G3	13.9	0.208	330-660
G4	19.8	0.297	360-870
G5	19.9	0.298	520-1 000
G6	8.8	0.132	330-545
G7	9.5	0.142	387-676
G8	8.0	0.120	580-828
G10	39.2	0.588	330-1 000
G11	29.5	0.442	360-960
G12	83.7	1.256	520-1 000

基金项目：国家自然科学基金（U2031144）资助.

收稿日期：2023-05-06；修订日期：2023-05-26

作者简介：宋得阳，男，硕士，主要研究天文技术与方法. Email：dysong2020@niaot.ac.cn.

* 通信作者：王 靓，男，副研究员，主要研究天体物理. Email：lwang@niaot.ac.cn.

程序引入了优化抽谱方法，自动屏蔽了宇宙线，提高了最终数据产品的精度。用户可以使用该程序快速实现批量数据处理，大大减轻了用户的工作量，有效提升了工作效率。

1 系统整体框架

1.1 环境配置

本程序基于 Python3 编程语言编写，依赖一些开源和免费的第 3 方软件包，包括 Astropy(用于 FITS 文件的读取与存储)，Numpy，Scipy(用于科学计算)和 Matplotlib(可视化)。本软件在 www.zenodo.org/record/7866030#.ZEi9uM5ByUk 以开源形式发布。

1.2 程序结构与流程

程序主要包括数据读取、数据合并和光谱抽取 3 部分。其中数据读取包括扫描生成观测日志、读取各个步骤数据。数据合并包括本底合并、平场合并与改正以及定标灯红蓝端合并。光谱抽取包括本底改正、平场改正、目标星定位、图像畸变改正、背景改正、波长定标和一维谱抽取等步骤。具体流程如图 1。用户只需要输入原始数据文件所在的路径，程序自动执行必要的步骤并存储各步骤产生的中间结果，以便后续使用，对关键步骤进行可视化并得到最终数据产品。

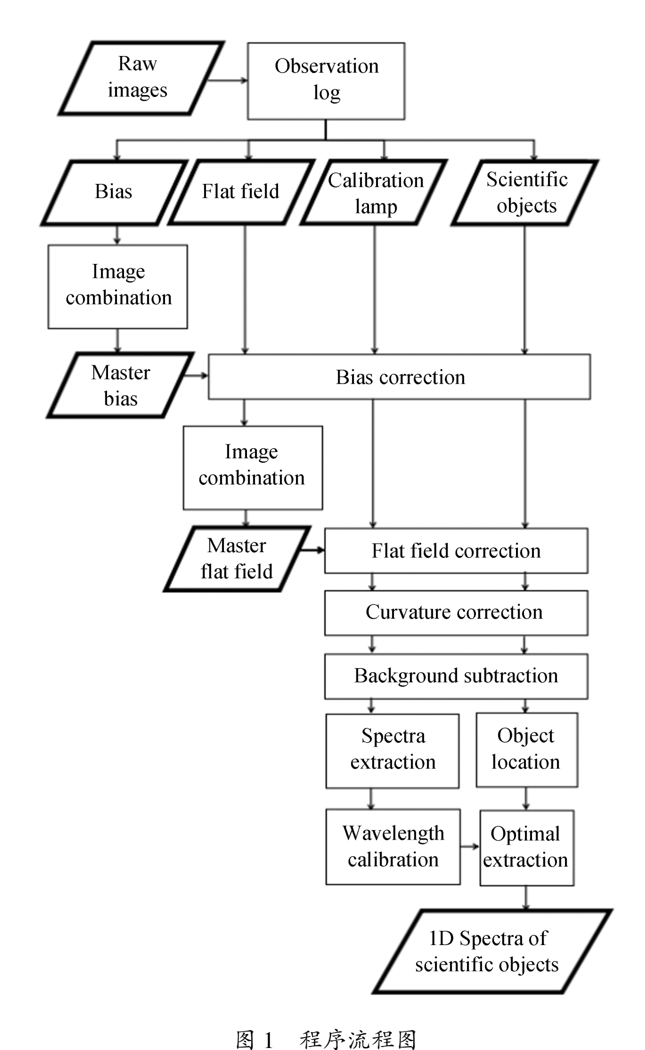


图 1 程序流程图
Fig. 1 Pipeline flow chart

其中本底(bias)是指没有光子输入时，CCD 像素在通电时(零秒曝)的读出值。平场(flat)是指输入光源为均匀光的情况下 CCD 获得的图像，反映了 CCD 的各个像素对光子的响应差异。定标灯是指使用空心阴极射线灯的光谱，包含波长已知的发射谱线，用于标定 CCD 图像上波长与像素之间的对应关系。科学目标是光谱仪的使用者拍摄的天体目标，例如彗星、恒星和星系等。

1.3 程序使用

本文介绍的程序是针对 BFOSC 的长缝光谱设计，用于将光谱仪获取的原始观测数据(通常为 FITS 格式的光谱图像)转换为可供科学分析的一维光谱，即目标天体的波长与流量的对应关系，最终输出数据以 FITS 格式保存。用户需要在操作系统的终端运行本程序，并输入原始数据文件和观测日志所在的路径，程序自动执行必要的步骤，存储中间结果并进行可视化，得到最终数据产品，即目标天体的一维谱。这些结果和中间过程文件都存储在指定的计算机目录中。此外，用户也可以更改运行模式，手动对目标星进行标记，程序会根据标记位置寻找目标星准确轮廓。

2 实际处理问题与解决方法

2.1 扫描生成观测日志

2.16 m 望远镜的观测者在观测同时经常会记录电子版的观测日志，但是为了避免格式不一致，

程序会扫描原始数据所在的计算机目录，读取其中 FITS 格式的文件，根据每个文件的头部(Header)信息，结合用户记录的观测日志，重新生成可以机读的格式化观测日志，包含文件名、图像类型、观测目标源的名称、曝光时间、观测开始时间等信息。

2.2　本底改正

通常情况下，观测者在观测开始前或结束后在 CCD 相机快门关闭的情况下进行多次零秒曝光，获得一组本底图像，通过图像合并获得本底的数值并抑制 CCD 读出噪声带来的误差。这个步骤提供了两种不同的合并模式，即均值模式和中值模式，通常情况下默认为均值模式。本底合并后生成本底图像，并将合并后的本底图像以 FITS 格式单独存储为一个文件。之后对需要操作的数据直接扣除保存的本底合并数据即完成本底改正。

2.3　目标源定位

本文介绍的程序采用以下方法确定目标天体的光谱在 CCD 图像上的位置：从光谱图像的左侧开始，沿 x 轴方向每隔 50 个像素取图像的截面最高处作为初始位置，取上下各 20 像素范围内的数据，用高斯函数进行拟合，获取纵向(y 方向)的中心位置与半高全宽，并将中心位置当作新的初始位置。重复上述过程，直到拟合参数收敛。接下来对所有 y 方向的中心位置与横坐标(x 值)的关系进行三阶多项式函数拟合，作为目标源光谱的中心位置曲线。光谱图像如图 2。此外程序还为用户提供了手动选取目标源的模式，在该模式下用户可以在弹出的图形界面上用鼠标标注目标源的大致位置，程序在其标记的位置附近按上述流程寻找目标天体的光谱。

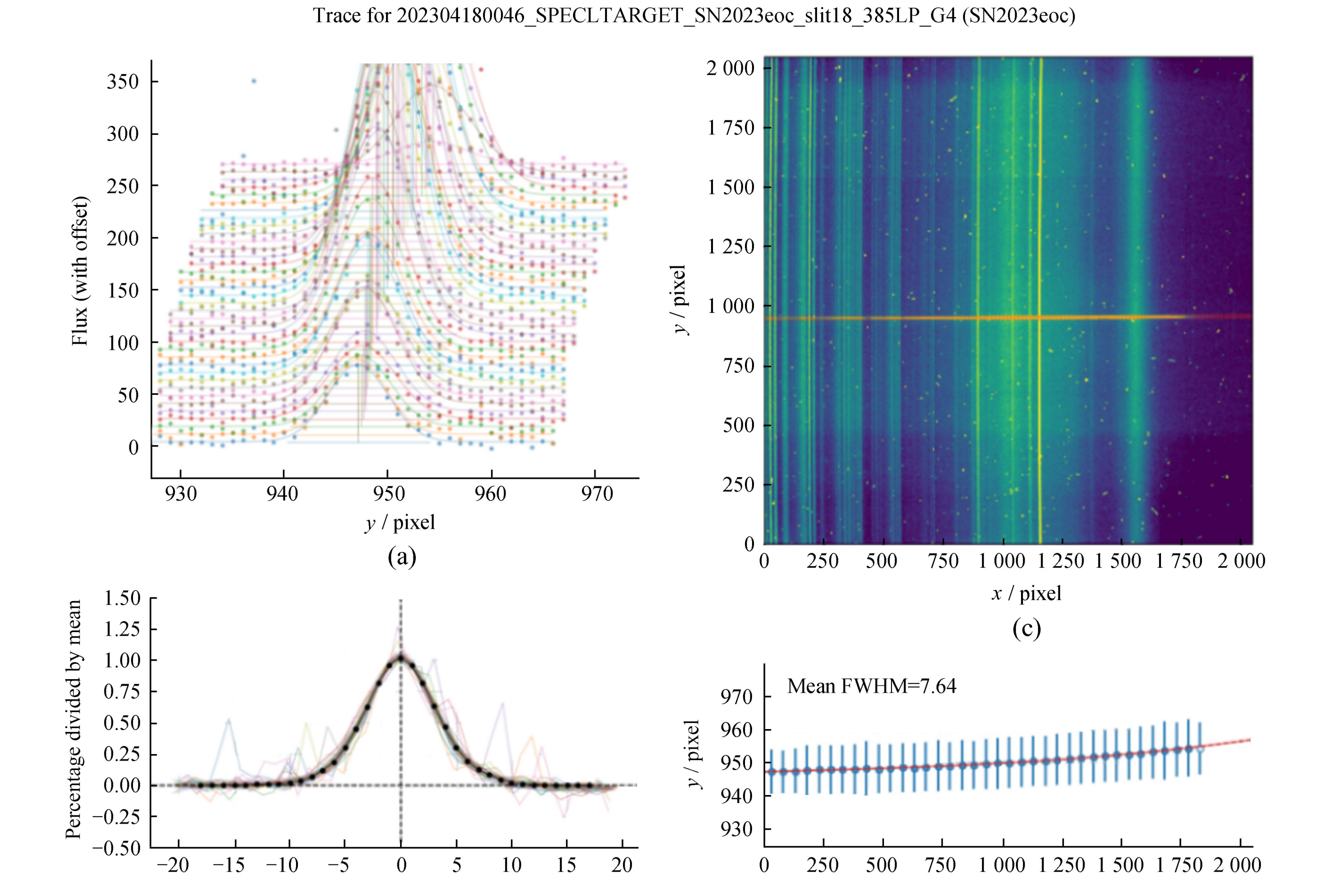


图 2　(a)中心位置与半高全宽图；(b)平均轮廓图；(c)真实目标源位置图；(d)轮廓中心拟合图

Fig. 2　(a) The central positions and full-width half-maximum (FWHM) of the spectral profiles, (b) the average profile is plotted in the lower left panel with black curve. (c) the target image, and (d) the positions and FWHMs of the center fittings

2.4 平场改正

由于 CCD 的各个像素对光子的响应略有差异，即便输入光源为均匀光，CCD 获得的图像也是不均匀的。在拍摄过程中一般进行多次曝光，将获得的多幅平场图像进行叠加，以获得较高的信噪比并抑制读出噪声。然后使用三次样条插值补齐数据中的坏像元数据，之后使用 Savitzky-Golay 滤波过程[8]获得主平场图像。每一幅目标天体的光谱图像需要除以主平场图像进行平场改正。

2.5 视场光学畸变改正

由于 CCD 相机的光学畸变，CCD 图像上狭缝所成的像存在一定程度的弯曲。在波长定标灯的光谱图像上选取沿着 y 轴中间位置的一维谱作为基准光谱，在 y 轴的上下两个方向，每间隔 200 个像素进行一次同样的一维谱提取操作，将得到的光谱与基准光谱计算交叉相关函数

$$r_k^{xy} = \frac{\frac{1}{n}\sum_{t=1}^{n-k}(y_t - \bar{y})(x_{t+k} - \bar{x})}{\sqrt{SD_x SD_y}} , \tag{1}$$

其中，x 和 y 分别为待改正的光谱与基准光谱；n 为光谱取点个数；k 为相对像素差值；SD_x 和 SD_y 分别为两光谱的标准差。交叉相关函数最大值对应的像素之差即为实际的像素偏移量。根据各行的像素差获得改正函数，并生成改正图像，如图 3，不同颜色的线代表 y 方向不同位置的光谱。经过视场畸变改正后能有效提高一维光谱的波长定标精度，避免系统误差。

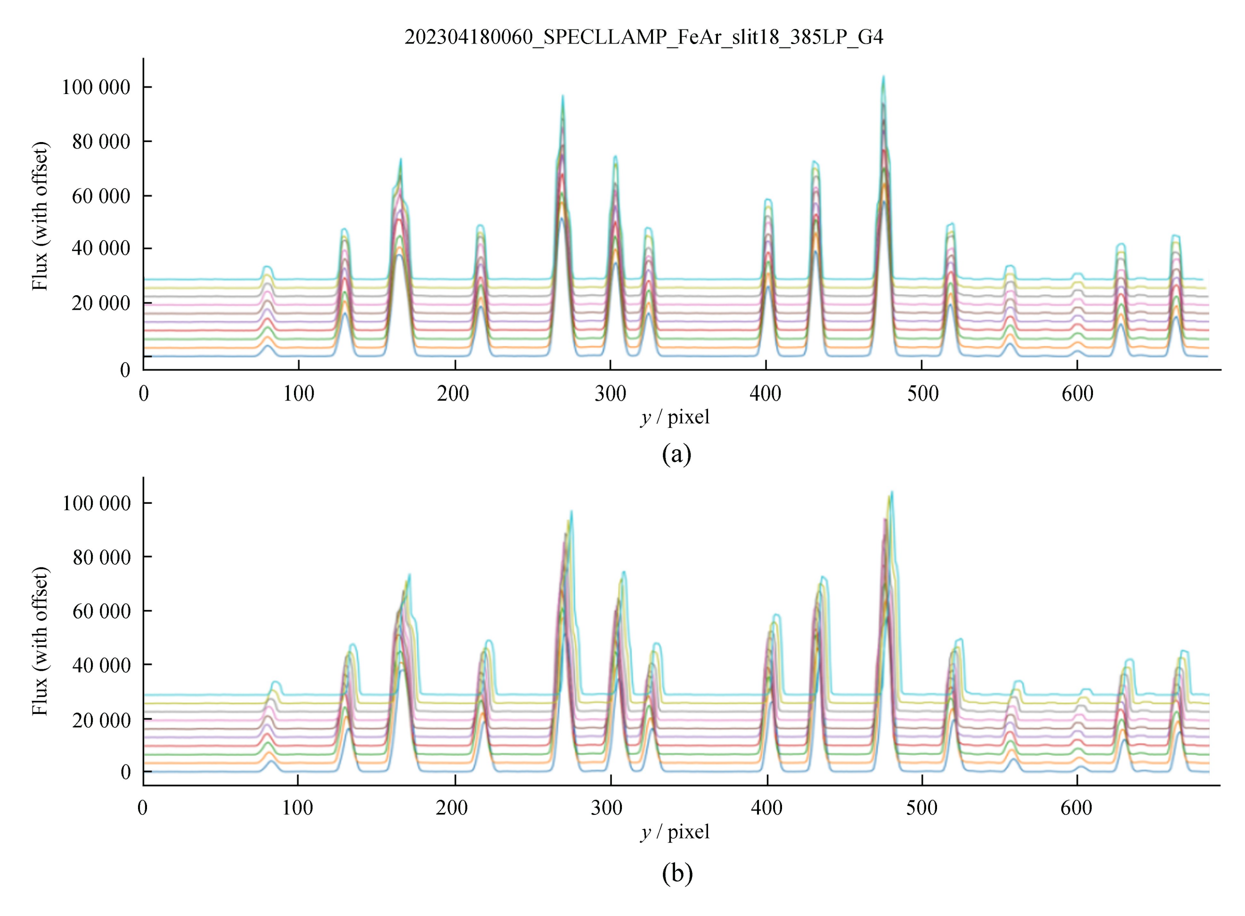


图 3 (a)畸变改正后的 FeAr 定标灯谱的一维谱图像；(b)畸变改正前的一维谱图像

Fig. 3 (a) The FeAr spectra at different positions along the slit direction after the distortion correction; (b) before the distortion correction

2.6 背景改正

实际观测过程中，我们获得的光谱不仅来自于目标源(图 4(a))，也有天光背景以及仪器内部的

杂散光。目标星的光谱也包含这些成分，因此为了获取目标源的真实流量，我们需要对图像数据进行背景改正。

程序一般会控制目标星使其落在沿 y 轴中间附近的位置。程序自动选取 CCD 图像上距离目标星足够远、CCD 像素一致性比较好的行(例如图 4(b)中第 340~500 行以及 1 600~1 820 行)作为测定背景光大小的区域。在选取过程中，程序会计算该区域每一行沿着 x 轴(即色散方向)的平均值和标准差，对于强度超过均值以上 3 倍标准差的行，可能包含目标源附近其他天体的光谱，这种情况在密集星场中尤为常见。程序在计算背景值时对这些行进行了屏蔽，以免获得比真实值显著偏高的背景流量。目标源光谱附近的像素值减去背景光谱，即为目标天体的光谱。

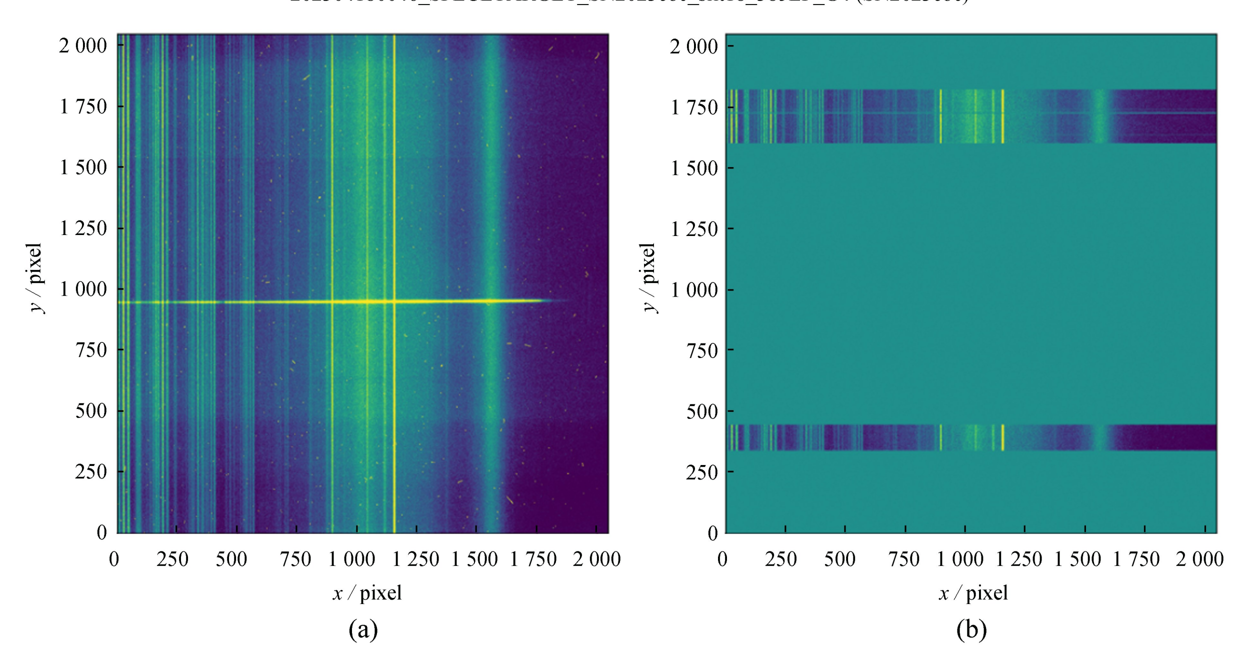


图 4　(a)目标源天体的原始光谱图像；(b)本例中测定天空背景采用的像素区域

Fig. 4　(a) The original 2D image of the target source; (b) the CCD image areas that are used to determine the background level

2.7　波长定标

波长定标的目的是在 CCD 图像上建立像素位置与波长的映射关系。本文介绍的程序包含自动对波长进行标定，具体处理流程如下：(1)读取该模式下提供的已经定标完成的模板光谱数据、模板特征发射线波长与像素位置对应数据；其次对要处理的光谱数据与模板光谱数据进行交叉相关函数计算，获得两个光谱数据位置的偏移量，将读入的模板像素位置数据增加偏移量获得新的像素位置数据。(2)依次取新的像素位置左右各 10 个像素的一维光谱，用广义高斯函数 $A\exp[-(a-c)/\alpha]^{\beta}+B$ 进行拟合，获得峰值位置对应的像素数值。结果如图 5 和图 6，其中每幅图中的圆点代表一维光谱数据，实线代表拟合曲线，虚线代表拟合得到的谱线中心位置。(3)用五阶多项式函数进行拟合，获得波长与像素位置的对应关系函数，并保存全部波长与像素位置数据组，画出波长定标函数图像，如图 7。

此外，由于 BFOSC 采用的定标灯本身的特性，其红端存在大量明亮的 Ar 元素的发射线，而蓝端以 Fe 线为主，它们的强度差异较大，如果曝光时间过长会导致红端饱和，而曝光时间过短会导致蓝端谱线没有足够的强度。因此，我们通常拍摄一次长曝光和一次短曝光，分别作为蓝端和红端的波长定标谱。在本文的例子中，红蓝两端的波长定标图像曝光时间分别为 30 s 和 300 s。因此，在进行定标之前本方法对两组光谱数据进行截取之后合并处理生成一个新的定标光谱图像，处理过程中计算交叉相关函数时对红端和蓝端分别求解计算。

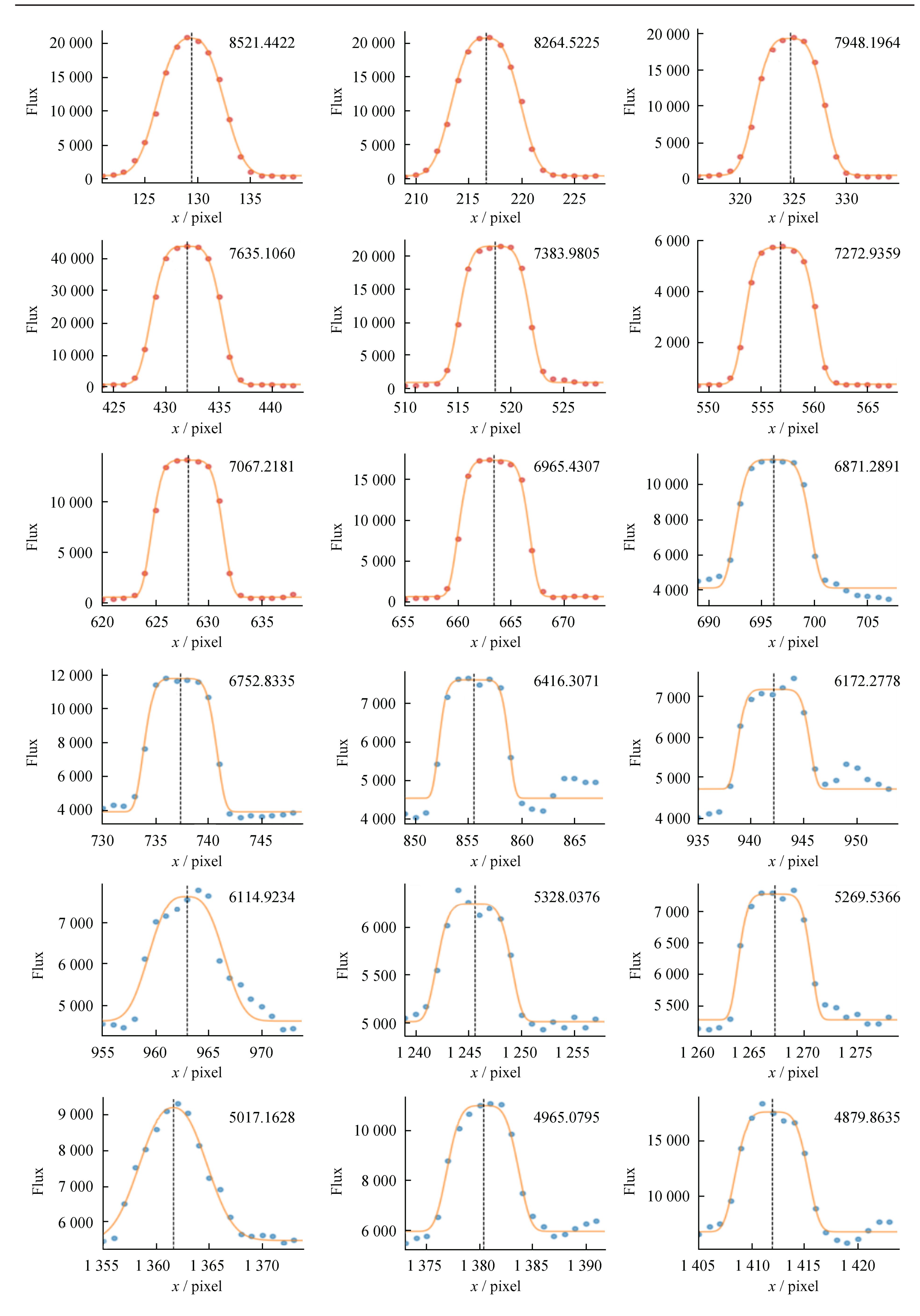


图 5 用于波长证认的 FeAr 谱线峰值位置拟合

Fig. 5 Peak positions of the identified FeAr lines

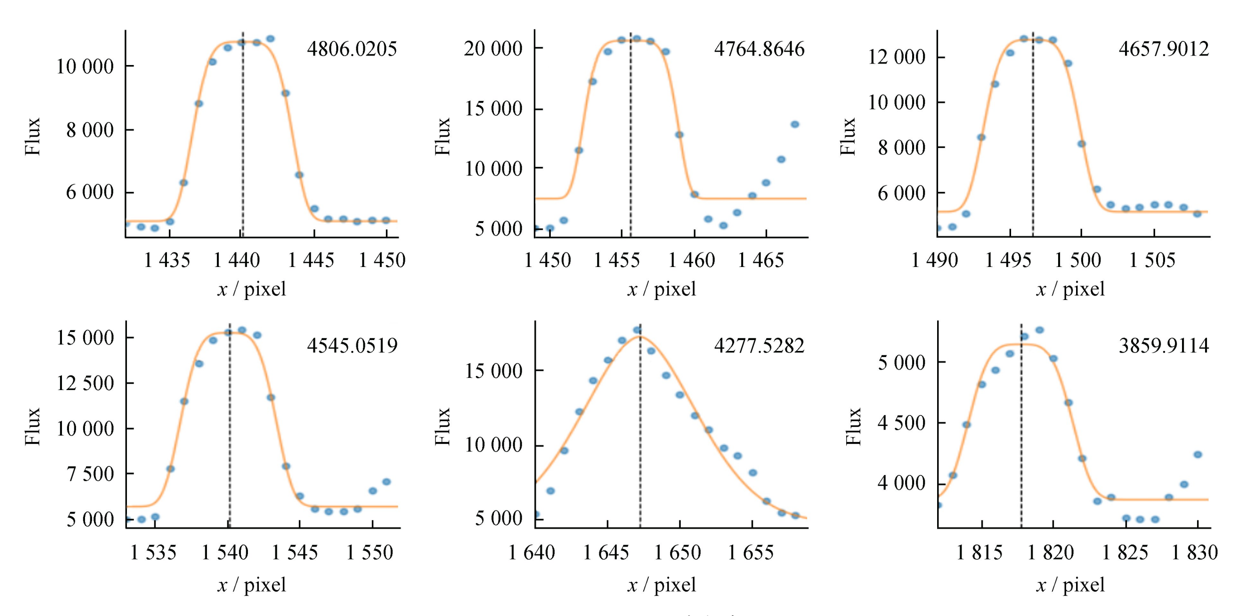


图 6　用于波长证认的 FeAr 谱线峰值位置拟合

Fig. 6　Peak positions of the identified FeAr lines

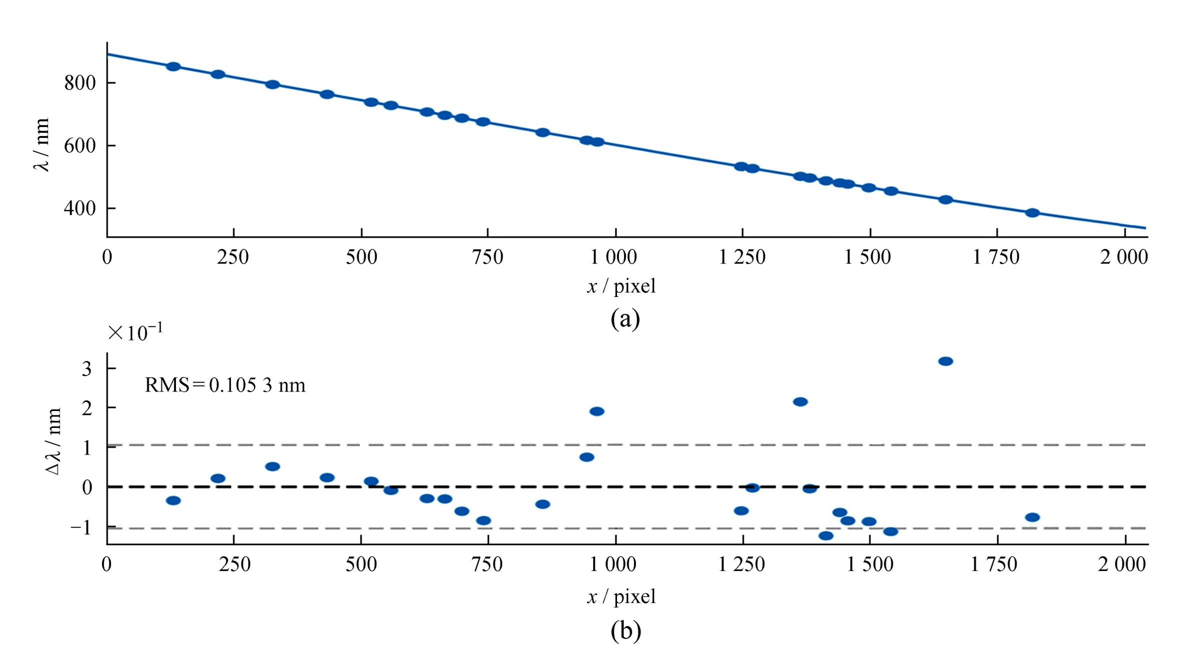


图 7　(a) 多项式拟合得到的波长(y)与像素(x)的关系，蓝色点代表定标谱线；(b) 拟合残差

Fig. 7　(a) The relation of fitted wavelength v. s. pixel numbers, where blue dots represent the fitting emission lines; (b) the residuals

2. 8　优化抽谱

优化抽谱方法采用加权求和，保证具有足够流量的情况下包含尽量少的噪声像素，对暗弱源尤其有效，能显著提高信噪比[9]。优化抽谱过程主要分为两步：(1) 获得目标星的精确空间轮廓，确定每个像素的加权值；(2) 根据空间轮廓的形状，排除明显高于轮廓的像素点，即扣除宇宙线。在我们开发的光谱数据处理程序中，第 1 步首先默认空间轮廓宽度为 36 个像素，轮廓方向每个像素具有相同的权重值，然后计算光谱数据方差，根据方差改正加权值与空间轮廓，重复此过程迭代 2~3 次，最终获得精确的空间轮廓。第 2 步，先根据上一步得到的轮廓提取科学目标的一维谱，并与普通抽谱方法得到的一维谱计算差值，对其中大于 5 倍标准偏差的像素点进行屏蔽，我们认为这些像素是受到宇宙线照射的像素。重复上述步骤直至收敛，至此我们获得了屏蔽宇宙线后的科学目标一维谱。

3 结果比较

为了验证程序的可靠性，我们使用 IRAF 对超新星 SN2023eoc 的同一组光谱数据进行了处理并绘制了一维谱图像，如图 8，其中，黄色线是本软件自动化处理获得的结果，蓝色线是 IRAF 处理结果。由图 8 可以看出，两种处理方法结果基本一致，并且本软件有效自动屏蔽了宇宙线，在流量比较低的蓝端，信噪比明显优于 IRAF 的处理结果。IRAF 需要在处理过程中对每条光谱手动屏蔽宇宙线，相比之下本文方法节省了处理时间且精度更高。

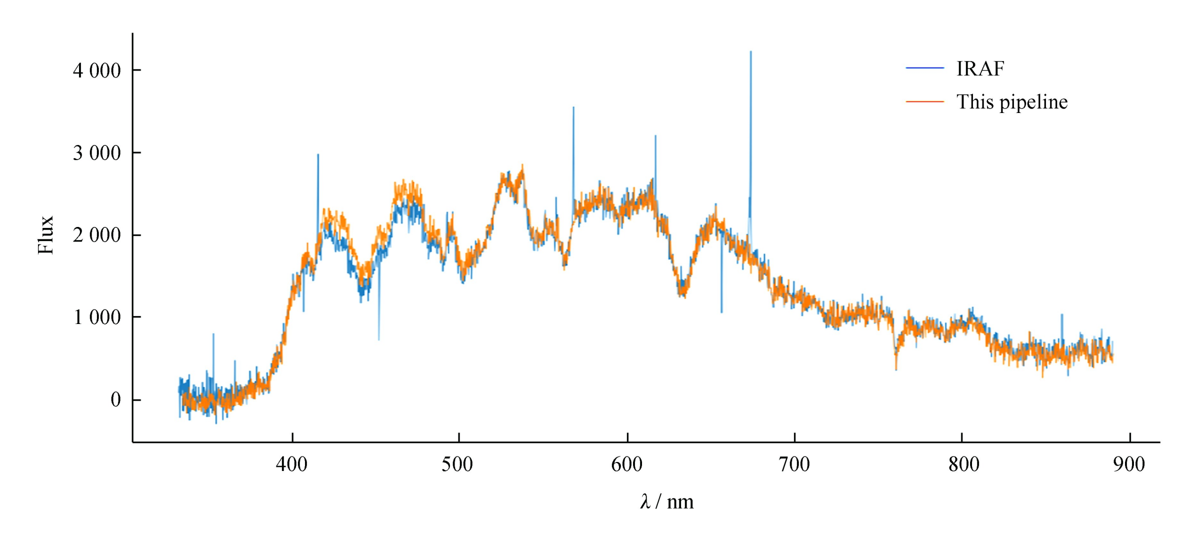


图 8　结果比较。蓝色曲线为 IRAF 处理结果，黄色曲线为本软件处理结果

Fig. 8　Comparison of results. The blue curve is the IRAF result, and the yellow curve is the result of this pipeline

4 总　结

本文系统介绍了一款针对 2.16 m 望远镜 BFOSC 光谱仪的长缝光谱数据处理软件，主要流程包括生成日志、本底和平场的合并与改正、目标星定位、视场畸变改正、背景改正、波长定标、优化抽谱。本文的数据处理软件使用方便，相比 IRAF 需要手动处理，本软件只需要用户输入原始数据所在的路径即可，处理过程透明可视，每个步骤完成都会生成相应的图片并保存数据以便重复使用，工作效率高，能批量化自动处理光谱数据且自动扣除宇宙线数据。与 IRAF 处理结果相比效果更好，具有较高的实用价值。

致谢：感谢兴隆 2.16 m 望远镜全体工作人员的支持。

参考文献：

[1] FAN Z, WANG H, JIANG X, et al. The Xinglong 2.16-m telescope: current instruments and scientific projects [J]. Publications of the Astronomical Society of the Pacific, 2016, 128(969): 115005.

[2] 黄磷，吴宏，李宏斌. BFOSC 使用说明 [EB/OL]. (2012-04-15) [2023-05-06]. http://astrocloud.china-vo.org/s/static/u/doc/BFOSCmanualv2.1_chinese.doc.
HUANG L, WU H, LI H B. BFOSC operating manual [EB/OL]. (2012-04-15) [2023-05-06]. http://astrocloud.china-vo.org/s/static/u/doc/BFOSCmanualv2.1_chinese.doc.

[3] 兴隆观测站，中国科学院国家天文台. 终端介绍 [EB/OL]. (2012-04-15) [2023-05-

06］. http://www.xinglong-naoc.cn/html/gcyq/216/detail-18.html.
Xinglong Observatory, National Astronomical Observatories, Chinese Academy of Sciences. About the telescope and instrument ［EB/OL］.（2012-04-15）［2023-05-06］. http://www.xinglong-naoc.cn/html/gcyq/216/detail-18.html.
［4］ 郑捷，江林巧. 光变曲线自动抽取程序［J］. 天文研究与技术，2023，20(1)：83-89.
ZHENG J, JIANG L Q. An automatic light-curve extraction pipeline ［J］. Astronomical Research & Technology, 2023, 20(1): 83-89.
［5］ PERLEY D A. Fully automated reduction of longslit spectroscopy with the low resolution imaging spectrometer at the Keck Observatory ［J］. Publications of the Astronomical Society of the Pacific, 2019, 131(1002): 084503.
［6］ TODY D. The IRAF data reduction and analysis system ［C］// Proceedings of SPIE. 1986.
［7］ 张艳，林众，郭迪福. IRAF 软件中 DAOPHOT 包的应用［J］. 天文研究与技术，2011，8(4)：410-416.
ZHANG Y, LIN Z, GUO D F. An application of the DAOPHOT package in the IRAF ［J］. Astronomical Research & Technology, 2011, 8(4): 410-416.
［8］ SAVITZKY A, GOLAY M. Smoothing and differentiation of data by simplified least squares procedures ［J］. Analytical Chemistry, 1964, 36(8): 1627-1639.
［9］ HORNE K. An optimal extraction algorithm for CCD spectroscopy ［J］. Publications of the Astronomical Society of the Pacific, 1986, 98(604): 609.

BFOSC Long-slit Spectrum Automatic Extraction Pipeline

Song Deyang[1,2,3], Zhang Junbo[4,5], Wang Liang[1,2,3*]

(1. Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences, Nanjing 210042, China, Email: lwang@niaot.ac.cn; 2. CAS Key Laboratory of Astronomical Optics & Technology, Nanjing Institute of Astronomical Optics & Technology, Nanjing 210042, China; 3. University of Chinese Academy of Sciences, Beijing 100049, China; 4. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; 5. Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China)

Abstract: In astronomical observation, the long-slit spectrometer is a kind of spectral observation equipment commonly used. The long-slit spectral mode of the Beijing Faint Object Spectrograph and Camera (BFOSC) mounted on the 2.16 m telescope of the Xinglong Observatory, National Astronomical Observatories, Chinese Academy of Sciences, is widely used, but it is not equipped with a specific data processing software. After observation, users usually run IRAF software interactively to obtain the final data product. This paper introduces an automatic data processing software for BFOSC long slit spectrum. It includes bias and flat-field correction, background subtraction, spectrum extraction and wavelength calibration. We describe the processing method and the flow of the program, and present the results of processing with the observed data, and compare them with the IRAF results. Our software realizes automatic operation on the basis of completing all the necessary steps of spectrum processing, and adopts the optimal extraction method, which automatically deducts most cosmic rays and ensures the high precision of output results.

Key words: long-slit spectrum; optimized extraction method; automatic pipeline